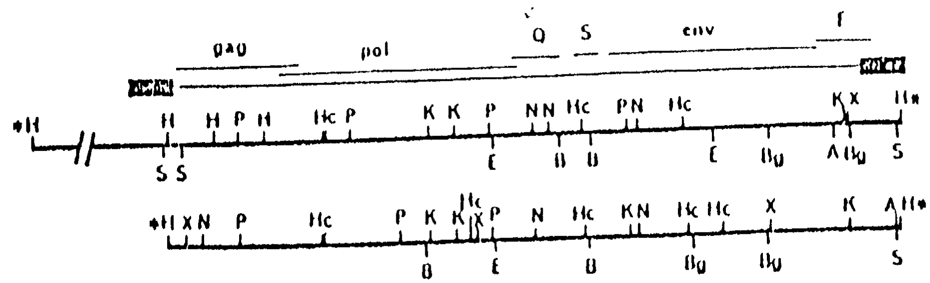
gag
pol
Q
S
env
F
LAV eli
LAV mal

FIG. 1A

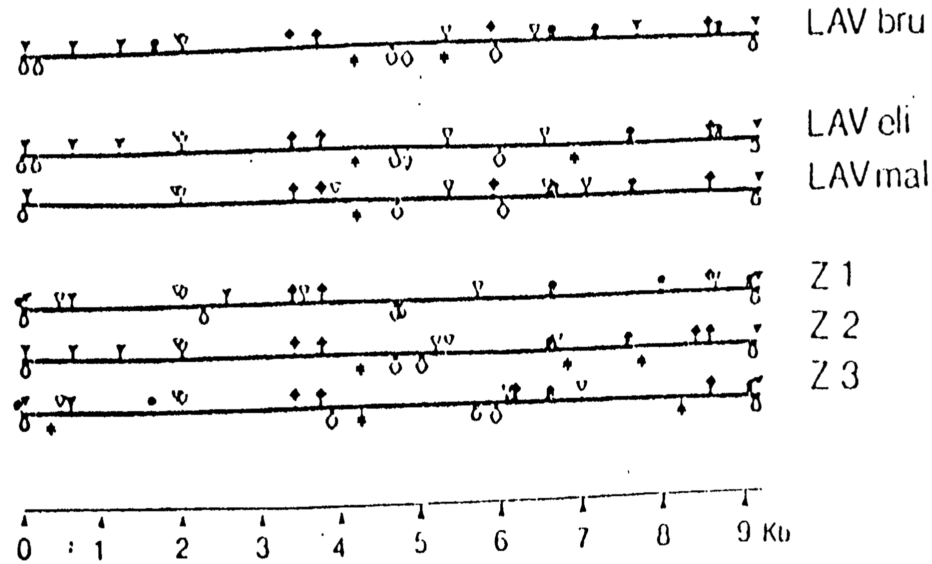
LAV bru
LAV eli
LAV mal
Z 1
Z 2
Z 3
0 1 2 3 4 5 6 7 8 9 Kb

FIG. 1B

FIG. 2

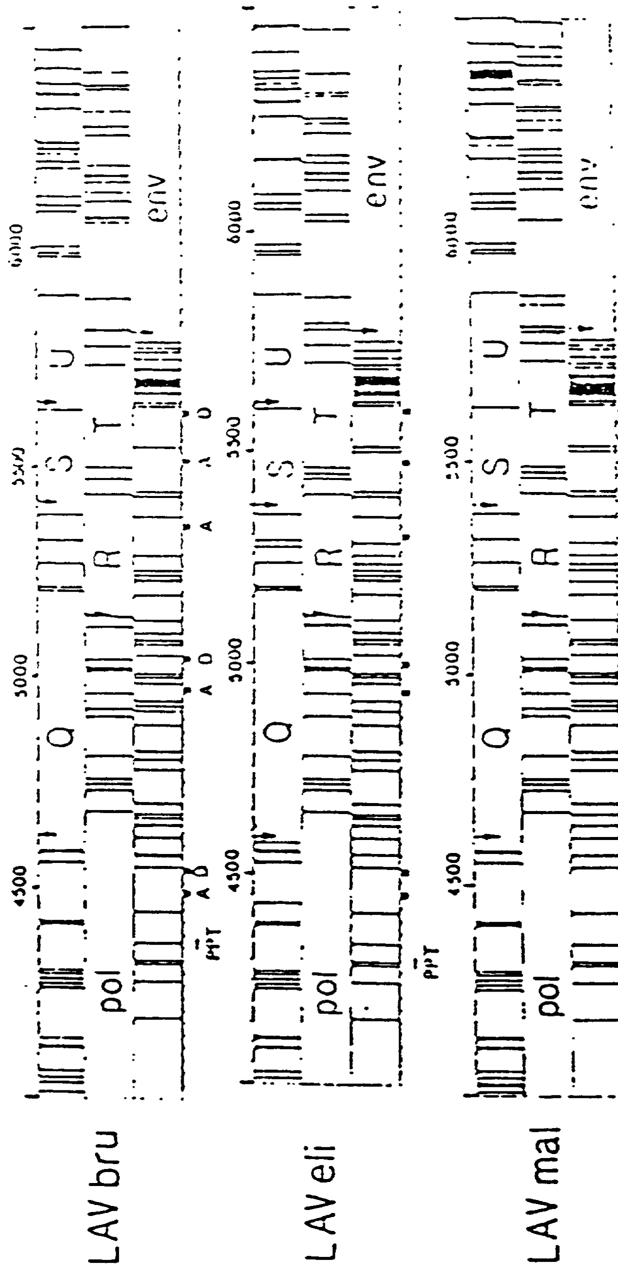

FIG. 3A

Coll
66

GAG

```
                10         20         30         40         50         60         70         80
LAV BRU MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI LGQLQPSLQT GSEELRSLYN
ARV 2                  K
LAV HAL             K  A               R    L          L        G   Q   HE   ST K      IK
LAV ELI             K  K               R                 Y L           K   I     AI     T

                90        100        110        120        130        140        150        160
LAV BRU TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA ------DTGH SSQVSQNYPI VQNIQGQMVH QAISPRTLNA
ARV 2                DV         E                    ----AAG  H               L
LAV HAL              DV         .    I     RQ T      AQQAAAA KH   S              A    I
LAV ELI          K G DV         E H                  ------   H N                L

               170        180        190        200        210        220        230        240
LAV BRU WVKVVEEKAF SPEVIPMFSA LSEGATPQDL NTMLNTVGGH QAAMQMLKET INEEAAEWDR VHPVHAGPIA PGQMREPRGS
ARV 2
LAV HAL .   I                            H   I             D          D            P
LAV ELI     I                                                                L

               250        260        270        280        290        300        310        320
LAV BRU DIAGTTSTLQ EQIGWMTNNP PIPVGEIYKR WIILGLNKIV RMYSPTSILD IRQGPKEPFR DYVDRFYKTL RAEQASQEVK
ARV 2                                                                                       D
LAV HAL                   S          D                    V                      F          T
LAV ELI               A   S                     V         V                                 D

               330        340        350        360        370        380        390        400
LAV BRU NWMTETLLVQ NANPDCKTIL KALGPAATLE EMMTACQGVG GPGHKARVLA EAMSQVTNS- ATIMMQRGNF RNQRKIVKCF
ARV 2                                                                P-  H               T
LAV HAL                           G                 S             A   T  A          KG - KI
LAV ELI                           Q                 S             A   V T A         KGP   I

               410        420        430        440        450        460        470        480
LAV BRU NCGKEGHIAR NCRAPRKKGC WKCGKEGHQM KDCTERQANF LGKIWPSYKG RPGNFLQSRP EPTAPPFLQS RPEPTAPPEE
ARV 2            K             R  R                                              ---- --------
LAV HAL        L                                       H                         ---- ---------A
LAV ELI          K             R       L            K    H                       ---- ---------A

               490        500        510
LAV BRU SFRSGVETTT PSQKQEPIDK ELYPLTSLRS LFGNDPSSQ
ARV 2      F E K
LAV HAL   GF E IK-       QK       A  K       QL
LAV ELI   GF E I -       QK          K        L
```

FIG. 33

Central region : Q

```
                   10         20         30         40         50         60         70         80
LAV BRU    MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSCKARGWFY RHHYESPHPR ISSEVHIPLC DARLVITTYW GLHTGERDWH
ARV 2                                      I K  K          T    V            K                E
LAV MAL                          H           K  KH         R  K V                VR        Q   E
LAV ELI                        K             K  HK            K          E        K           E

                   90        100        110        120        130        140        150        160
LAV BRU    LGQGVSIEWR KKRYSTQVDP ELADQLIHLY YFDCFSDSAI RKALLGHIVS PRCEYQAGHN KVCSLQYLAL AALITPKKIK
ARV 2           A        K                                                            
LAV MAL    N          Q       L  G     H      L    KH I  YR                              T
LAV ELI                R         D             E    Q I          D                 T   A   TR
                                 G     H       E      I  D                          T   A  Q

                  170        180        190
LAV BRU    PPLPSVTKLT EDRWNKPQKT KGHRGSHTMN GH
ARV 2           K
LAV MAL         R          Q
LAV ELI         R          Q  R
```

R

```
                   10         20         30         40         50         60         70         80
LAV BRU    MEQAPEDQGP QREPHNEWTL ELLEELKNEA VRHFPRIWLH GLGQHIYETY GDTWAGVEAI IRILQQLLFI HFRIGCRHSR
ARV 2                      Y                R          P   S   Y
LAV MAL        A                            Q              S                E         S        Q
LAV ELI        A           Y    A           S              S                V                  Q
                                                                                               Q

                   90
LAV BRU    IGVTQQRRAR -NGASRS
ARV 2       II         R
LAV MAL     I R        -  S
LAV ELI     IIR        -  S
```

S (tat)

```
                   10         20         30         40         50         60         70
LAV BRU    MEPVDPRLEP WKHPGSQPKT ACTTCYCKKC CFHCQVCFTT KALGISYGRK KRRQRRRPPQ GSQTHQVSLS KQ
ARV 2            N              R   HH           YA   R  G                A   D    A
LAV MAL      D   N       N      R  P HK       Y  H   I  G                     H A   DP P L
LAV ELI      D   N       N      R  P HK H     Y  P   LH G               G      G A   PIP
```

FIG. 3C

POL

```
                 10         20         30         40         50         60         70         80
LAV BRU  FPREDLAFLQ GKAREFSSEQ TRANSPTFSS EQTRANSPTR RELQVWGRDN NSLSEAGADR QCTVSFNFPQ ITLWQRPLVT
ARV 2                                 --- ---------         GE
LAV HAL      H   P        P           --- ---------S    R   G - KT   T  E    I   S           V
LAV ELI      H   P     G L PK         --- ---------S    R     -  P  KT  E                      A

                 90        100        110        120        130        140        150        160
LAV BRU  IKIGGQLKEA LLDTGADDTV LEENSLPGRW KPKMIGGIGG FIKVRQYDQI LIEICGHKAI GTVLVGPTPV NIIGRNLLTQ
ARV 2     R                        N   K                        PV
LAV HAL  VRV                      IN   K                              K      I              H
LAV ELI                            H   K                        P     Q

                170        180        190        200        210        220        230        240
LAV BRU  IGCTLNFPIS PIETVPVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK EGKISKIGPE NPYNTPVFAI KKKDSTKWRK
ARV 2
LAV HAL                              R                T   KD        L
LAV ELI                                               T    D         R           I

                250        260        270        280        290        300        310        320
LAV BRU  LVDFRELNKR TQDFWEVQLG IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR YQYNVLPQGW
ARV 2                                                        K
LAV HAL    H
LAV ELI                                                                  S

                330        340        350        360        370        380        390        400
LAV BRU  KGSPAIFQSS MTKILEPFRK QNPDIVIYQY MDDLYVGSDL EIGQHRTKIE ELRQHLLRWG LTTPDKKHQK EPPFLWMGYE
ARV 2                                                                      F
LAV HAL                      T K  E                                E   K   F
LAV ELI                           EH                            K  E       F R

                410        420        430        440        450        460        470        480
LAV BRU  LHPDKWTVQP IVLPEKDSWT VNDIQKLVGK LNWASQIYPG IKVRQLCKLL RGTKALTEVI PLTEEAELEL AENREILKEP
ARV 2                H                            A     K
LAV HAL              Q  D E                             K         A    DIV    A
LAV ELI           S  K    E         H  ER

                490        500        510        520        530        540        550        560
LAV BRU  VHGVYYDPSK DLIAEIQKQG QGQWTYQIYQ EPFKNLKTGK YARTRGAHTN DVKQLTEAVQ KITTESIVIW GKTPKFKLPI
ARV 2      E          V                                 H                   VS          I
LAV HAL                                    QY           IKS                  AQ             K
LAV ELI                        H                        H            A     K S        K     R
```

```
                  570        580        590        600        610        620        630        640
LAV BRU    QKETWETWWT EYWQATWIPE WEFVNTPPLV KLWYQLEKEP IVGAETFYVD GAASRETKLG KAGYVTKRGR QKVVTLTDTT
ARV 2            A  H                                                N             D        SIA
LAV HAL          A                                 T                 N    K        D        S  E
LAV ELI             A                                   I            N             D        P

                  650        660        670        680        690        700        710        720
LAV BRU    NQKTELQAIH LALQDSGLEV NIVTDSQYAL GIIQAQPDKS ESELVNQIIE QLIKKEKVYL AWVPAHKGIG GNEQVDKLVS
ARV 2                                                       S
LAV HAL                      S                            I          Q D     S
LAV ELI             N

                  730        740        750        760        770        780        790        800
LAV BRU    AGIRKVLFLD GIDKAQDEHE KYHSNWRAMA SDFNLPPVVA KEIVASCDKC QLKGEAMHGQ VDCSPGIWQL DCTHLEGKVI
ARV 2               N       E                                                                I
LAV HAL    S                E                      I                                            I
LAV ELI    Q                E       N

                  810        820        830        840        850        860        870        880
LAV BRU    LVAVHVASGY IEAEVIPAET GQETAYFLLK LAGRWPVKTI HTDNGSNFTS TTVKAACWWA GIKQEFGIPY NPQSQGVVES
ARV 2
LAV HAL    I                            I           VV            AA         N
LAV ELI                                             VV            AA

                  890        900        910        920        930        940        950        960
LAV BRU    MNKELKKIIG QVRDQAEHLK TAVQMAVFIH NFKRKGGIGG YSAGERIVDI IATDIQTKEL QKQITKIQNF RVYYRDSRDP
ARV 2        N                                                                                NK
LAV HAL                  E                                    I M                              H
LAV ELI                                         NR            I                  I

                  970        980        990       1000       1010
LAV BRU    LWKGPAKLLW KGEGAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED
ARV 2
LAV HAL    I                                               G G
LAV ELI    I                     K           V
```

FIG. 3D

ENV

```
                                       SP                           OMP
                10         20         30 ●       40         50        ● 60         70        ● 80
LAV BRU   MRVK---EKY QHLWRWGWKW GTMLLGILMI CSATEKLWVT VYYGVPVWKE ATTTLFCASD AKAYDTEVHN VWATHACVPT
ARV 2      K  GTRRN             --- - L    M                                K
LAV HAL      REIQRN    NW       --- -N     N  T   IA D                        S E       I
LAV ELI     ARGIEKN[C] NW K     --- -I        T    ADN                        S E   A   I

                90        100        110        120 ●      130 ●      140        150        160
LAV BRU   DPNPQEVVLV NVTENFNMWK NDMVEQMHED IISLWDQSLK PCVKLTPLCV SLKCTDL-GN ATNTNSSNTN SSSGEMMME-
ARV 2              C .            N     Q                      T N    - K         --- ----KMKL I
LAV HAL         IE E       G      N                            T N  NVN T  V GTNAGS  RTNA LK  I
LAV ELI         IA E              N                            T N S E--L  RN GTNG NV TTLEKG----

            ●170        180        190        200        210 ●      220        230 ●      240
LAV BRU   KGEIKNCSFN ISTSIRGKVQ KEYAFFYKLD IIPIDNDTTS -----YTLTS CNTSVITQAC PKVSFEPIPI HYCAPAGFAI
ARV 2                T   D I    N L RK    VV     AS T THYTN K IN  R                     T
LAV HAL   -  V       TPVGSD R   -  T  N   LVQ  DSDN  ----S K IN                T D
LAV ELI   ---N       VT VLKD K  QV  L  K     V    SST -NSTN K IN    A

      ●    250 ●      260 ●      270        280        290        300        310 ●      320
LAV BRU   LKCNNKTFNG TGPCTNVSTV QCTHGIRPVV STQLLLNGSL AEEEVVIRSA NFTDNAKTII VQLNQSVEIN CTRPNNNTRK
ARV 2                K                 I                       D    N            E  A
LAV HAL      D K       EI K          K                    IN    E   L  T N      ET T     G    K
LAV ELI     RD K                                          I     E   L N   N    AN  E  K T  A  YQ   Q

                330        340       ●350        360        370        380        390      ● 400
LAV BRU   SIRIQRGPGR AFVTIGK-IG NMRQAHCNIS RAKWNATLKQ IASKLREQFG NNKT-IIFKQ SSGGDPEIVT HSFNCGGEFF
ARV 2       Y --       N T RI   DI K          Q  N  E   VK             - V N                M     K
LAV HAL   C NF--    Q  LY T I-V DI K Y T N ETE DK   Q   V V  GSLL- -   K   NS         T       R
LAV ELI   RTP -- L Q  SLY TKS-RS IIG          Q SK  Q   V R  GTLL- -   I K   P        T

      ●    410        420        430 ●      440        450        460 ●      470        480
LAV BRU   YCNSTQLFNS TWFNSTWSTE GSNNTEGSDT ITLPCRIKQF INMWQEVGKA MYAPPISGQI RCSSNITGLL LTRDGGNN--
ARV 2        T     N  -----RLN NTEG K N   I         I                    G  S                T -V
LAV HAL      TSK       Q NGARL- -  S  STGS           I     KT           A  V  N L       I      NSSD
LAV ELI      TSG       NI A NNI TES NSTNTN   Q      I   K VAGR-   I     ERN  L                  I --

                490        500        510        520       530|      540 TMP     550        560
LAV BRU   NNGSEIFRPG GGDMRDNWRS ELYKYKVVKI EPLGVAPTKA KRRVVQREKR AVGI-GALFL GFLGAAGSTM GARSMTLTVQ
ARV 2     T DT V                         I        I                  V  N              V L
LAV HAL    SDN TL             I          R                     E     I L-  N              A L
LAV ELI    STN T                         Q          K          E     I I-  N                V
```

FIG. 3E

FIG. 3F

```
                 570        580        590        600        610        620        630        640
LAV BRU   ARQLLSGIVQ QQNHLLRAIE AQQHLLQLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KLICTTAVPW NASWSNKSLE
ARV 2                                               V        R
LAV HAL                                             V        Q   R   H       H    P     S    K  D
LAV ELI       H                                                             H    H     S    K  N

                 650        660        670        680        690        700        710        720
LAV BRU   QIWNNMTWME WDREINNYTS LIHSLIEESQ NQQEKNEQLL LELDKWASLW NWFNITNWLW YIKIFIMIVG GLVGLRIVFA
ARV 2     D  D     Q  E   D   K T YT L                              S
LAV HAL   C        Q  EK  S   C I YN       I      K                 S SK      R   IV     I    I
LAV ELI   E  Q        E   D   G   Y        T      K                 S  Q            I    I

                 730        740        750        760        770        780        790        800
LAV BRU   VLSIVNRVRQ GYSPLSFQTH LPTPRGP-DR PEGIEEEGGE RDRDRSIRLV NGSLALIWDD LRSLCLFSYH RLRDLLLIVT
ARV 2                         R   V    -    D               V    D F     E           K         AA
LAV HAL      L             L  L        P              QG G         IS         H                A
LAV ELI      L                L   A    -      T        G    V  L   FS                       I  AV

                 810        820        830        840        850        860        870
LAV BRU   RIVELLGRRG WEALKYWWNL LQYWSQELKN SAVSLLNATA IAVAEGTDRV IEVVQGACRA IRHIPRRIRQ GLERILL
ARV 2      T  I  H           S      I          W         T          A R Y    L  H          L
LAV HAL                    L        G        I    T        Q       IG RYG    L          F  A
LAV ELI               DI   L            R    S  FD I               II R     VLN            S

                  10         20         30         40         50         60         70         80
LAV BRU  .MGGKWSKSSV VGWPTVRERM R----RAEPA ADGVGAASR- ----DLEKHG AITSSNTAAT NAACAWLEAQ EE-EEVGFPV
ARV 2            R H G  SAI      RAEP            V -  ----                    D          -
LAV HAL            I     KI   I  ---- TP T LT    V QD AVSQ  D C   AA  SP  N   S --- PP   L
LAV ELI            I     AI   I  ---- TH         V -  ----               S    D          SD

                  90        100        110        120        130        140        150        160
LAV BRU   TPQVPLRPMT YKAAVDLSHF LKEKGGLEGL IHSQRRQDIL DLWIYHTQGY FPDWQNYTPG PGVRYPLTFG WCYKLVPVEP
ARV 2     K              L I                W     E                           I          F
LAV HAL   K            GF              D    VV PK E      V                    I F        I    HS
LAV ELI   K            EL                   W KK  E      V H   I              I           E     D

                 170        180        190        200        210
LAV BRU   DKVEEANKGE NTSLLHPVSL HGMDDPEREV LEWRFDSRLA FHHVARELHP EYFKNC
ARV 2     E       E  K    H        E A K    V     K          H     Y D
LAV HAL   EE      E    HC   I Q    E A      K K   S   LR R   Q     Y D
LAV ELI   QE   DTE    TN    ICQ    E    Q   K   H      L K   H    FY -
```

FIG. 4A

A LAVbru vs.

	GAG		POL		ENV Total		ENV OMP		ENV TMP	
HTLV-3 USA	512 0/0	0.8	1015 0/0	1.3	856 5/0	1.4	507 5/0	1.6	349 0/0	1.1
ARV-2 USA	502 12/2	3.4	1003 12/0	3.1	855 17/11	13.0	505 17/10	14.3	350 0/1	11.2
LAVeli Zaire	500 13/1	9.8	1002 13/0	5.5	853 22/14	20.7	504 22/14	25.3	349 0/0	13.8
LAVmal Zaire	505 14/7	12.0	1002 13/0	7.7	859 13/11	21.7	509 13/10	26.4	350 0/1	14.9

B LAVeli vs.

	GAG		POL		ENV Total		ENV OMP		ENV TMP	
LAVmal	505 1/6	10.8	1002 0/0	8.4	859 13/11	19.8	509 8/13	23.6	350 0/1	14.3

FIG. 4B

	orf F		central region					
			orf Q		orf R		orf S	
A LAVbru vs.								
HTLV-3 USA	206 0/0	1.5	192 0/0	0	,	nd	80 0/0	2.5
ARV-2 USA	210 0/4	12.6	192 0/0	10.0	97 0/1	9.4	81 0/1	15.0
LAVeli Zaire	206 1/1	19.4	192 0/0	10.4	96 0/0	11.5	80 0/0	27.5
LAVmal Zaire	209 2/5	27.0	192 0/0	12.6	96 0/0	10.4	80 0/0	23.8
B LAVeli vs.								
LAVmal	209 3/6	22.5	192 0/0	12.0	96 0/0	6.3	80 0/0	11.3

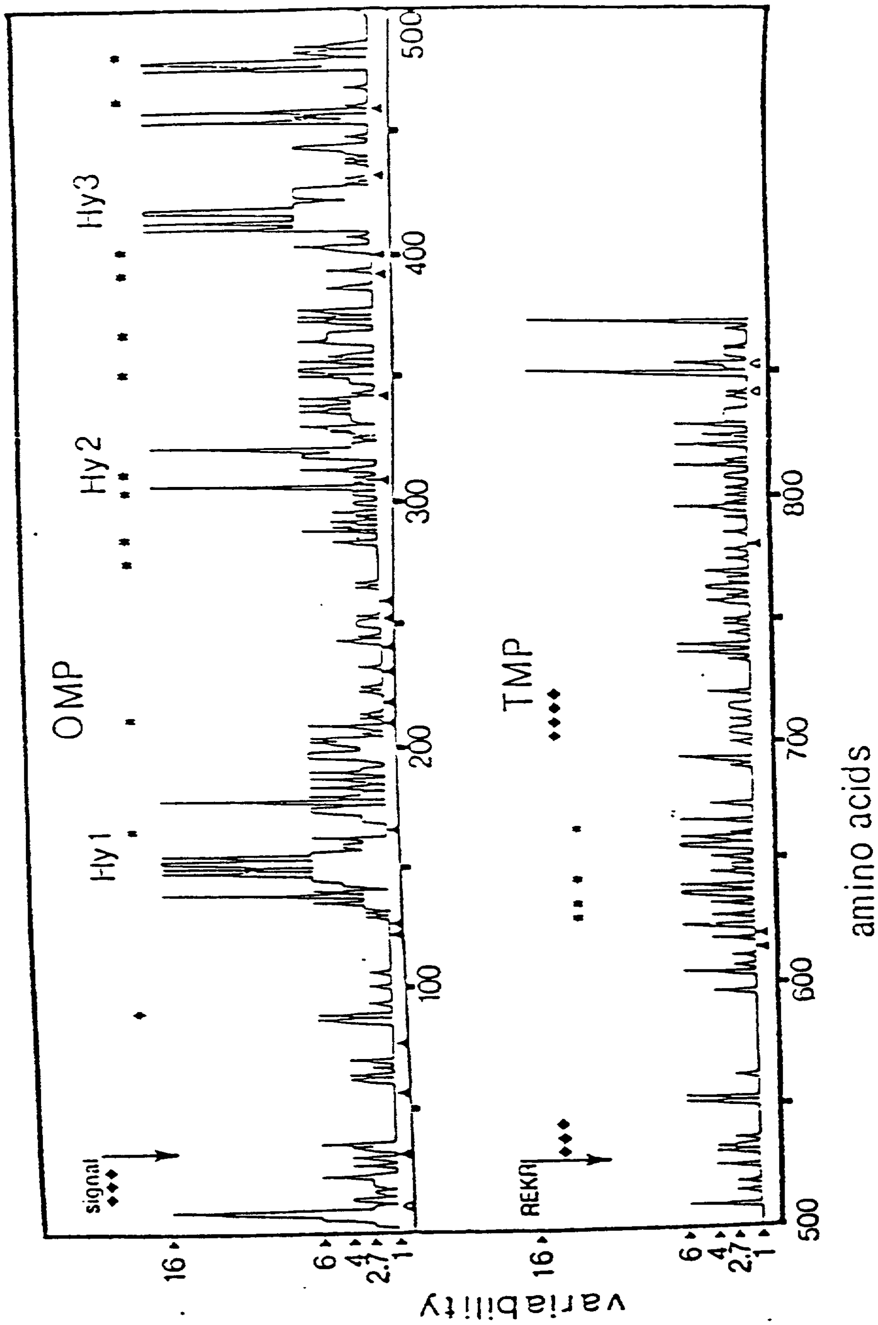

FIG. 5

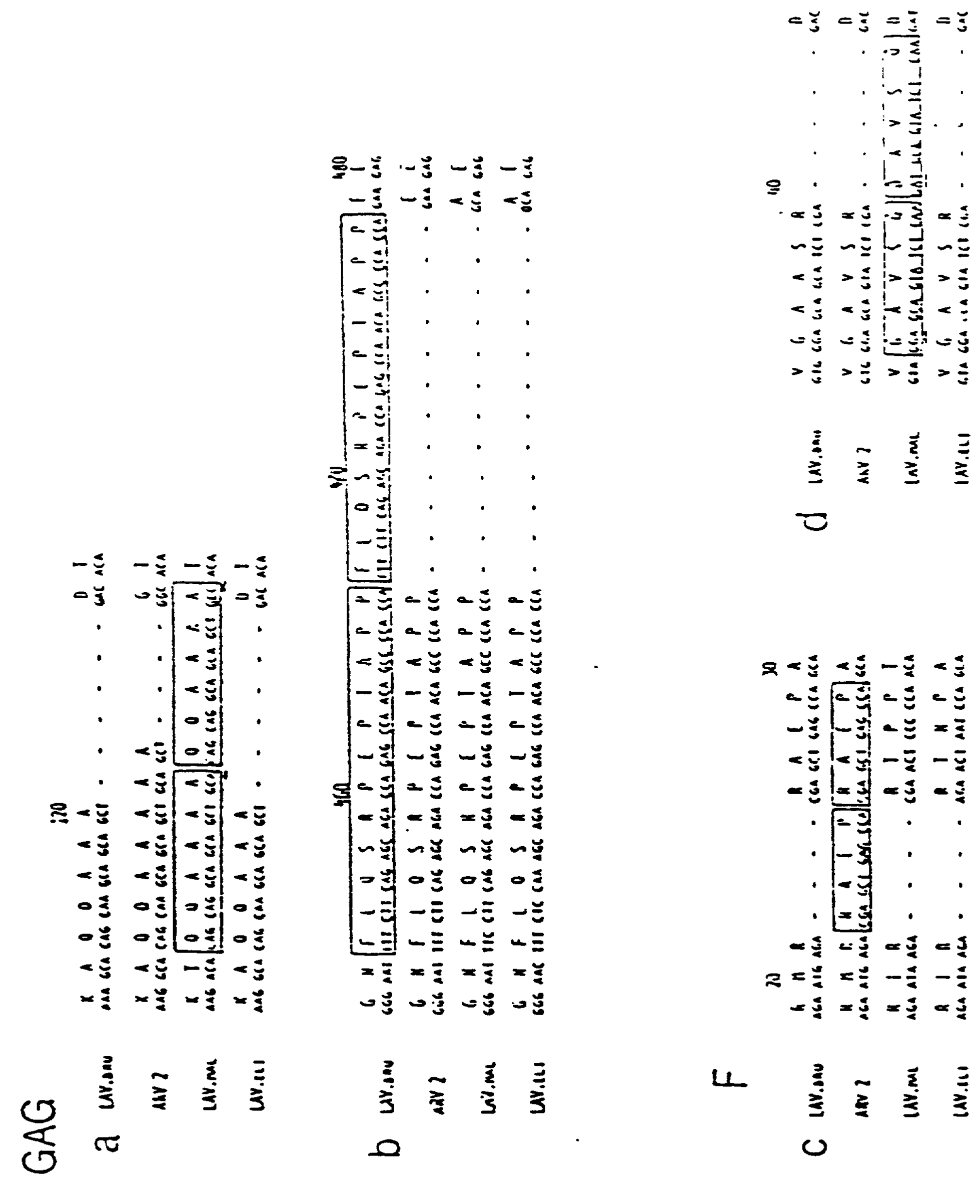

FIG. 6A

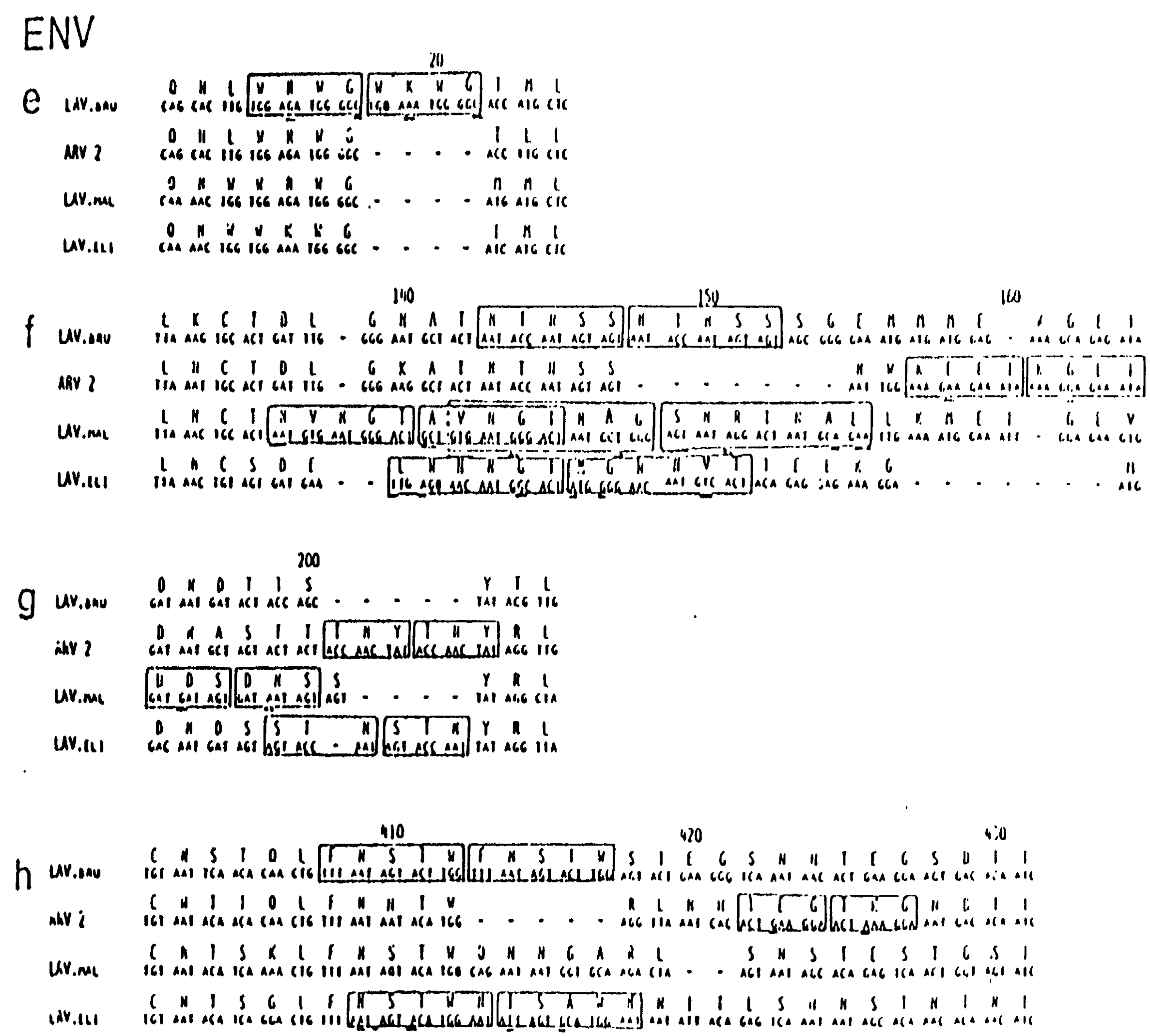
ENV
e
f
g
h
LAV.BRU
ARV 2
LAV.MAL
LAV.ELI
20
140
150
160
200
410
420
430

FIG. 6B

FIG. 7A

LAV.ELI

```
R
GGTCTCTCTGGTTAGACCAGATTTGAGCCTGGGAGCTCTCTGGCTAGCTAGGGAACCCAC
                                       R<-->.U5
TGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGT
                                         100
GTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTAGTCAGAGTGGAAAATCTCTAGCA
U5<
GTGGCGCCCGAACAGGGACCTGAAAGCGAAAGTAGAACCAGAGGAGCTCTCTCGACGCAG
             200
GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCAGCGACTGGTGAGTACGCT
                                       GAG                 300
                                       MetGlyAlaArgAlaSerValLeuSer
AAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAA

 GlyGlyLysLeuAspLysTrpGluLysIleArgLeuArgProGlyGlyLysLysLysTyr
GCGGGGGAAAATTAGATAAATGGGAAAAAATTCGGTTACGGCCAGGAGGAAAGAAAAAAT
                                         400
 ArgLeuLysHisIleValTrpAlaSerArgGluLeuGluArgTyrAlaLeuAsnProGly
ATAGACTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATATGCACTTAATCCTG

 LeuLeuGluThrSerGluGlyCysLysGlnIleIleGlyGlnLeuGlnProAlaIleGln
GCCTTTTAGAAACATCAGAAGGCTGTAAACAAATAATAGGGCAGCTACAACCAGCTATTC
             500
 ThrGlyThrGluGluLeuArgSerLeuTyrAsnThrValAlaThrLeuTyrCysValHis
AGACAGGAACAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTAC
                                                           600
 LysGlyIleAspValLysAspThrLysGluAlaLeuGluLysMetGluGluGluGlnAsn
ATAAAGGAATAGATGTAAAAGACACCAAGGAAGCTTTAGAAAAGATGGAGGAAGAGCAAA

 LysSerLysLysLysAlaGlnGlnAlaAlaAlaAspThrGlyAsnAsnSerGlnValSer
ACAAAAGTAAGAAAAAAGGCACAGCAAGCAGCAGCTGACACAGGAAACAACAGCCAGGTCA
                                         700
 GlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerPro
GCCAAAATTATCCTATAGTGCAGAACCTACAGGGGCAAATGGTACATCAGGCCATATCAC

 ArgThrLeuAsnAlaTrpValLysValIleGluGluLysAlaPheSerProGluValIle
CTAGAACTTTGAACGCATGGGTAAAAGTAATAGAAGAAAAGGCTTTCAGCCCAGAAGTAA
             800
 ProMetPheSerAlaLeuSerGluGlyAlaThrProGlnAspLeuAsnThrMetLeuAsn
TACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCATGCTAA
                                                           900
 ThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAla
ACACAGTGGGGGGACATCAAGCAGCCATGCAAATGCTAAAAGAGACCATCAATGAAGAAG

 AlaGluTrpAspArgLeuHisProValHisAlaGlyProIleAlaProGlyGlnMetArg
CTGCAGAATGGGATAGGTTACATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGA
                                        1000
 GluProArgGlySerAspIleAlaGlyThrThrSerThrLeuGlnGluGlnIleAlaTrp
GAGAACCAAGGGGAAGTGATATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGCAT

 MetThrSerAsnProProIleProValGlyGluIleTyrLysArgTrpIleIleValGly
GGATGACAAGTAACCCACCTATCCCAGTAGGAGAAATCTATAAAAGATGGATAATTGTGG
             1100
 LeuAsnLysIleValArgMetTyrSerProValSerIleLeuAspIleArgGlnGlyPro
GATTAAATAAAATAGTAAGAATGTATAGCCCTGTCAGCATTTTGGACATAAGACAGGGAC
                                                          1200
```

FIG. 7B

LysGluProPheArgAspTyrValAspArgPheTyrLysThrLeuArgAlaGluGlnAla
CAAAGGAACCTTTTAGAGACTATGTACACCGGTTCTATAAAACTCTAAGAGCCGAGCAAG

SerGlnAspValLysAsnTrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAsp
CTTCACAGGATGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCAAACCCAG
1300

CysLysThrIleLeuLysAlaLeuGlyProGlnAlaThrLeuGluGluMetMetThrAla
ATTGCAAGACTATCTTAAAAGCATTGGGACCACAGGCTACACTAGAAGAAATGATGACAG

CysGlnGlyValGlyGlyProSerHisLysAlaArgValLeuAlaGluAlaMetSerGln
CATGTCAGGGAGTGGGGGGCCCAGCCATAAAGCAAGAGTTCTGGCTGAGGCAATGAGCC
1400

AlaThrAsnSerValThrThrAlaMetMetGlnArgGlyAsnPheLysGlyProArgLys
AAGCAACAAATTCAGTTACTACAGCAATGATGCAGAGAGGCAATTTTAAGGGCCCAAGAA
1500

IleIleLysCysPheAsnCysGlyLysGluGlyHisIleAlaLysAsnCysArgAlaPro
AAATTATTAAGTGTTTCAATTGTGGCAAAGAAGGGCACATAGCAAAAAATTGCAGGGCCC

ArgLysLysGlyCysTrpArgCysGlyLysGluGlyHisGlnLeuLysAspCysThrGlu
CTAGGAAAAAGGGCTGTTGGAGATGTGGAAAGGAAGGACACCAACTAAAAGATTGCACTG
1600

→ POL

PhePheArgGluAsnLeuAlaPheProGlnGlyLysAlaGlyGluLeu
ArgGlnAlaAsnPheLeuGlyArgIleTrpProSerHisLysGlyArgProGlyAsnPhe
AGAGACAGGCTAATTTTTTAGGGAGAATTTGGCCTTCCCACAAGGGAAGGCCGGGGAACT

SerProLysGlnThrArgAlaAsnSerProThrSerArgGluLeuArgValTrpGlyArg
LeuGlnSerArgProGluProThrAlaProProAlaGluSerPheGlyPheGlyGluGlu
TTCTCCAAAGCAGACCAGAGCCAACAGCCCCACCAGCAGAGAGCTTCGGGTTTGGGGAAG
1700

AspAsnProLeuSerLysThrGlyAlaGluArgGlnGlyThrValSerPheAsnPhePro
IleThrProSerGlnLysGlnGluGlnLysAspLysGluLeuTyrProLeuThrSerLeu
AGATAACCCCCTCTCAAAAACAGGAGCAGAAAGACAAGGAACTGTATCCTTTAACTTCCC
1800

GAG ←

GlnIleThrLeuTrpGlnArgProLeuValAlaIleLysIleGlyGlyGlnLeuLysGlu
LysSerLeuPheGlyAsnAspProLeuSerGln
TCAAATCACTCTTTGGCAACGACCCCTTGTCGCAATAAAAATAGGGGGACAGCTAAAGGA

AlaLeuLeuAspThrGlyAlaAspAspThrValLeuGluGluMetAsnLeuProGlyLys
AGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAA
1900

TrpLysProLysMetIleGlyGlyIleGlyGlyPheIleLysValArgGlnTyrAspGln
ATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCA

IleProIleGluIleCysGlyGlnLysAlaIleGlyThrValLeuValGlyProThrPro
AATACCCATAGAAATCTGTGGACAGAAAGCTATAGGTACAGTATTAGTAGGACCTACGCC
2000

ValAsnIleIleGlyArgAsnLeuLeuThrGlnIleGlyCysThrLeuAsnPheProIle
TGTCAACATAATCGGAAGAAATTTGTTGACCCAGATTGGCTGCACTTTAAATTTTCCAAT
2100

SerProIleGluThrValProValLysLeuLysProGlyMetAspGlyProLysValLys
TAGTCCTATTGAAACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAA

GlnTrpProLeuThrGluGluLysIleLysAlaLeuThrGluIleCysThrAspMetGlu
ACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAACAGAAATTTGTACAGATATGGA
2200

FIG. 7C

```
LysGluGlyLysIleSerArgIleGlyProGluAsnProTyrAsnThrProIlePheAla
AAAGGAAGGAAAAATTTCAAGAATTGGGCCTGAAAATCCATACAATACTCCAATATTTGC

IleLysLysLysAspSerThrLysTrpArgLysLeuValAspPheArgGluLeuAsnLys
CATAAAGAAAAAAGACAGTACCAAGTGGAGAAAATTAGTAGATTTCAGAGAACTTAATAA
            2300
ArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAlaGlyLeuLysLys
GAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCGCATCCTGCAGGGCTGAAAAA
                                                        2400
LysLysSerValThrValLeuAspValGlyAspAlaTyrPheSerValProLeuAspGlu
GAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGA

AspPheArgLysTyrThrAlaPheThrIleSerSerIleAsnAsnGluThrProGlyIle
AGATTTTAGGAAATATACCGCCTTTACCATATCTAGTATAAACAATGAGACACCAGGGAT
                                    2500
ArgTyrGlnTyrAsnValLeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSer
TAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCGGCAATATTCCAAAG

SerMetThrLysIleLeuGluProPheArgLysGlnAsnProGluMetValIleTyrGln
TAGCATGACAAAAATCTTAGAGCCCTTTAGAAAACAAAATCCAGAAATGGTTATCTATCA
            2600
TyrMetAspAspLeuTyrValGlySerAspLeuGluIleGlyGlnHisArgThrLysIle
ATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGGACAAAAAT
                                                        2700
GluLysLeuArgGluHisLeuLeuArgTrpGlyPheThrArgProAspLysLysHisGln
AGAGAAATTAAGAGAACATCTATTGAGGTGGGGATTTACCAGACCAGATAAAAAACATCA

LysGluProProPheLeuTrpMetGlyTyrGluLeuHisProAspLysTrpThrValGln
GAAAGAACCCCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACA
                                    2800
SerIleLysLeuProGluLysGluSerTrpThrValAsnAspIleGlnAsnLeuValGlu
GTCTATAAAACTGCCAGAAAAGGAGAGCTGGACTGTCAATGATATACAGAACTTAGTGGA

ArgLeuAsnTrpAlaSerGlnIleTyrProGlyIleLysValArgGlnLeuCysLysLeu
GAGATTAAACTGGGCAAGCCAGATTTATCCAGGAATTAAAGTAAGACAATTATGTAAACT
            2900
LeuArgGlyThrLysAlaLeuThrGluValIleProLeuThrGluGluAlaGluLeuGlu
CCTTAGGGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAATTAGA
                                                        3000
LeuAlaGluAsnArgGluIleLeuLysGluProValHisGlyValTyrTyrAspProSer
ACTGGCAGAAAACAGGGAAATTTTAAAAGAACCAGTACATGGAGTGTATTATGACCCATC

LysAspLeuIleAlaGluIleGlnLysGlnGlyHisGlyGlnTrpThrTyrGlnIleTyr
AAAAGACTTAATAGCAGAAATACAGAAACAAGGGCACGGCCAATGGACATACCAAATTTA
                                    3100
GlnGluProPheLysAsnLeuLysThrGlyLysTyrAlaArgMetArgGlyAlaHisThr
TCAAGAACCATTTAAAAATCTGAAAACAGGAAAGTATGCAAGAATGAGGGGTGCCCACAC

AsnAspValLysGlnLeuAlaGluAlaValGlnArgIleSerThrGluSerIleValIle
TAATGATGTAAAGCAATTAGCAGAGGCAGTGCAAAGAATATCCACAGAAAGCATAGTGAT
            3200
TrpGlyArgThrProLysPheArgLeuProIleGlnLysGluThrTrpGluThrTrpTrp
ATGGGGAAGGACTCCTAAATTTAGACTACCCATACAAAAGGAAACATGGGAAACATGGTG
                                                        3300
```

FIG.7D

```
AlaGluTyrTrpGlnAlaThrTrpIleProGluTrpGluPheValAsnThrProProLeu
GGCAGAGTATTGGCAAGCCACTTGGATTCCTGAGTGGAATTTGTCAATACCCCTCCTTT

ValLysLeuTrpTyrGlnLeuGluLysGluProIleIleGlyAlaGluThrPheTyrVal
AGTAAAATTATGGTACCAGTTAGAGAAGGAACCCATAATAGGAGCAGAAACTTTCTATGT
                                            3400
AspGlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGlyTyrValThrAspArgGly
AGATGGGGCAGCTAATAGAGAGACTAAATTAGGAAAAGCAGGATATGTTACTGACAGAGG

ArgGlnLysValValProLeuThrAspThrThrAsnGlnLysThrGluLeuGlnAlaIle
AAGACAGAAAGTTGTCCCTTTGACTGACACGACAAATCAGAAGACTGAGTTACAAGCAAT
              3500
AsnLeuAlaLeuGlnAspSerGlyLeuGluValAsnIleValThrAspSerGlnTyrAla
TAATCTAGCCTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGATTCACAATATGC
                                                           3600
LeuGlyIleIleGlnAlaGlnProAspLysSerGluSerGluLeuValAsnGlnIleIle
ATTAGGAATCATTCAAGCACAACCAGATAAGAGTGAATCAGAGTTAGTCAATCAAATAAT

GluGlnLeuIleLysLysGluLysValTyrLeuAlaTrpValProAlaHisLysGlyIle
AGAGCAGTTAATAAAAAAGGAAAAGGTTTACCTGGCATGGGTACCAGCACACAAAGGAAT
                                            3700
GlyGlyAsnGluGlnValAspLysLeuValSerGlnGlyIleArgLysValLeuPheLeu
TGGAGGAAATGAACAAGTAGATAAATTAGTCAGTCAAGGAATCAGGAAAGTACTATTTTT

AspGlyIleAspLysAlaGlnGluGluHisGluLysTyrHisAsnAsnTrpArgAlaMet
GGATGGAATAGATAAGGCTCAAGAAGAACATGAGAAATATCACAACAATTGGAGAGCAAT
              3800
AlaSerAspPheAsnLeuProProValValAlaLysGluIleValAlaSerCysAspLys
GGCTAGTGATTTTAACCTACCACCCGTGGTAGCAAAAGAAATAGTAGCTAGCTGTGATAA
                                                           3900
CysGlnLeuLysGlyGluAlaMetHisGlyGlnValAspCysSerProGlyIleTrpGln
ATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCA

LeuAspCysThrHisLeuGluGlyLysValIleLeuValAlaValHisValAlaSerGly
ATTAGATTGTACACACTTAGAAGGAAAAGTTATCCTGGTAGCAGTTCATGTAGCCAGTGG
                                            4000
TyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeu
CTATATAGAAGCAGAAGTTATTCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTT

LysLeuAlaGlyArgTrpProValLysValValHisThrAspAsnGlySerAsnPheThr
AAAATTAGCAGGAAGATGGCCAGTAAAAGTAGTACATACAGACAATGGCAGCAATTTCAC
              4100
SerAlaAlaValLysAlaAlaCysTrpTrpAlaGlyIleLysGlnGluPheGlyIlePro
CAGTGCTGCAGTTAAGGCCGCCTGTTGGTGGGCAGGTATCAAACAGGAATTTGGAATTCC
                                                           4200
TyrAsnProGlnSerGlnGlyValValGluSerMetAsnLysGluLeuLysLysIleIle
CTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTAT

GlyGlnValArgAspGlnAlaGluHisLeuLysThrAlaValGlnMetAlaValPheIle
AGGACAGGTAAGAGATCAAGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCAT
                                            4300
HisAsnPheLysArgArgArgGlyIleGlyGlyTyrSerAlaGlyGluArgIleIleAsp
CCACAATTTTAAAAGAAGAAGGGGGATTGGGGGATACAGTGCAGGGGAAAGAATAATAGA
```

FIG. 7E

```
IleIleAlaThrAspIleGlnThrLysGluLeuGlnLysGlnIleIleLysIleGlnAsn
CATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTATAAAAATTCAAAA
        4400

PheArgValTyrTyrArgAspSerArgAspProIleTrpLysGlyProAlaLysLeuLeu
TTTTCGGGTTTATTACAGAGACAGCAGAGATCCAATTTGGAAAGGACCAGCAAAGCTCCT
                                                        4500

TrpLysGlyGluGlyAlaValValIleGlnAspLysSerAspIleLysValValProArg
CTGGAAAGGTGAAGGGGCAGTAGTAATACAAGACAAGAGTGACATAAAGGTAGTACCAAG

                         Q
ArgLysValLysIleIleArgAspTyrGlyLysGlnMetAlaGlyAspAspCysValAla
                         MetGluAsnArgTrpGlnValMetIleValTrpGln
AAGAAAAGTAAAGATTATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGC
          POL                       4600

SerArgGlnAspGluAsp
ValAspArgMetArgIleLysThrTrpLysSerLeuValLysHisHisMetTyrValSer
AAGTAGACAGGATGAGGATTAAAACATGGAAAAGTTTAGTAAAACACCATATGTATGTTT

LysLysAlaAsnArgTrpPheTyrArgHisHisTyrGluSerProHisProLysIleSer
CAAAGAAAGCTAACAGATGGTTTTATAGACATCACTATGAAAGCCCCCACCCAAAAATAA
        4700

SerGluValHisIleProLeuGlyGluAlaArgLeuValIleLysThrTyrTrpGlyLeu
GTTCAGAAGTACACATCCCACTAGGAGAAGCTAGACTGGTAATAAAAACATATTGGGGTC
                                                        4800

HisThrGlyGluArgGluTrpHisLeuGlyGlnGlyValSerIleGluTrpArgLysArg
TGCATACAGGAGAAAGAGAATGGCATCTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAA

ArgTyrSerThrGlnValAspProGlyLeuAlaAspGlnLeuIleHisMetTyrTyrPhe
GGAGATATAGCACACAAGTAGACCCTGGCCTGGCAGACCAACTAATTCATATGTATTATT
                                    4900

AspCysPheSerGluSerAlaIleArgLysAlaIleLeuGlyAspIleValSerProArg
TTGATTGTTTTTCAGAATCTGCTATAAGAAAAGCCATATTAGGAGATATAGTTAGTCCTA

CysGluTyrGlnAlaGlyHisAsnLysValGlySerLeuGlnTyrLeuAlaLeuThrAla
GGTGTGAGTATCAAGCAGGACATAACAAGGTAGGATCCCTACAGTATTTGGCACTAACAG
        5000

LeuIleAlaProLysGlnIleLysProProLeuProSerValArgLysLeuThrGluAsp
CATTAATAGCACCAAAACAGATAAAGCCACCTTTGCCTAGTGTTAGGAAGCTAACAGAAG
    R                                                   5100

    MetGluGlnAlaProAlaAspGlnGlyProGlnArgGluProTyrAsnGluTrpAla
ArgTrpAsnLysProGlnGlnThrArgGlyHisArgGlySerHisThrMetAsnGlyHis
ATAGATGGAACAAGCCCCAGCAGACCAGGGGCCACAGAGGGAGCCATACAATGAATGGGC

Q
LeuGluLeuLeuGluGluLeuLysSerGluAlaValArgHisPheProArgIleTrpLeu
ATTAGAGCTTTTAGAGGAGCTTAAGAGTGAAGCTGTTAGACATTTTCCTAGGATATGGCT
                                    5200

HisSerLeuGlyGlnHisIleTyrGluThrTyrGlyAspThrTrpValGlyValGluAla
CCATAGCTTAGGACAACATATTTATGAAACTTATGGGGATACCTGGGTAGGAGTTGAAGC

IleIleArgIleLeuGlnGlnLeuLeuPheIleHisPheArgIleGlyCysGlnHisSer
TATAATAAGAATACTGCAACAATTACTGTTTATTCATTTCAGAATTGGGTGTCAACATAG
        5300                       S           R

ArgIleGlyIleIleArgGlnArgArgAlaArgAsnGlySerSerArgSer
                                    MetAspProValAspProAsnLeuGlu
CAGAATAGGCATTATTCGACAGAGAAGAGCAAGAAATGGATCCAGTAGATCCTAACCTAG
                                                        5400
```

FIG. 7F

```
 ProTrpAsnHisProGlySerGlnProArgThrProCysAsnLysCysHisCysLysLys
AGCCCTGGAACCATCCAGGAAGTCAGCCTAGGACTCCTTGTAACAAGTGTCATTGTAAAA

 CysCysTyrHisCysProValCysPheLeuAsnLysGlyLeuGlyIleSerTyrGlyArg
AGTGTTGCTATCATTGCCCAGTTTGCTTCTTAAACAAAGGCTTAGGCATCTCCTATGGCA
                                   5500
 LysLysArgArgGlnArgArgGlyProProGlnGlyGlyGlnAlaHisGlnValProIle
GGAAGAAGCGGAGACAGCGACGAGGACCTCCTCAAGGCGGTCAGGCTCATCAAGTTCCTA
     S
 ProLysGln
TACCAAAGCAGTAAGTAGTACATGTAATGCAACCTTTAGGGATAATAGCAATAGCAGCAT
              5600
TAGTAGTAGCAATAATACTAGCAATAGTTGTGTGGACCATAGTATTCATAGAATATAGAA
                                                         5700
GGATAAAAAAGCAAAGGAGAATAGACTGTTTACTTGATAGAATAACAGAAAGAGCAGAAG
         ENV
         MetArgAlaArgGlyIleGluArgAsnCysGlnAsnTrpTrpLysTrpGly
ACAGTGGCAATGAGAGCGAGGGGGATAGAGAGAAATTGTCAAAACTGGTGGAAATGGGGC
                                   5800
IleMetLeuLeuGlyIleLeuMetThrCysSerAlaAlaAspAsnLeuTrpValThrVal
ATCATGCTCCTTGGGATATTGATGACCTGTAGTGCTGCAGACAATCTGTGGGTCACAGTT

TyrTyrGlyValProValTrpLysGluAlaThrThrThrLeuPheCysAlaSerAspAla
TATTATGGGGTGCCTGTATGGAAGGAAGCAACCACCACTCTATTTTGTGCATCAGATGCT
              5900
LysSerTyrGluThrGluAlaHisAsnIleTrpAlaThrHisAlaCysValProThrAsp
AAATCATATGAAACAGAGGCACATAATATCTGGGCCACACATGCCTGTGTACCCACGGAC
                                                         6000
ProAsnProGlnGluIleAlaLeuGluAsnValThrGluAsnPheAsnMetTrpLysAsn
CCCAACCCACAAGAAATAGCACTGGAAAATGTGACAGAAAACTTTAACATGTGGAAAAAT

AsnMetValGluGlnMetHisGluAspIleIleSerLeuTrpAspGlnSerLeuLysPro
AACATGGTGGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAACCA
                                   6100
CysValLysLeuThrProLeuCysValThrLeuAsnCysSerAspGluLeuArgAsnAsn
TGTGTAAAATTAACCCCACTCTGTGTCACTTTAAACTGTAGTGATGAATTGAGGAACAAT

GlyThrMetGlyAsnAsnValThrThrGluGluLysGlyMetLysAsnCysSerPheAsn
GGCACTATGGGGAACAATGTCACTACAGAGGAGAAAGGAATGAAAAACTGCTCTTTCAAT
              6200
ValThrThrValLeuLysAspLysLysGlnGlnValTyrAlaLeuPheTyrArgLeuAsp
GTAACCACAGTACTAAAAGATAAGAAGCAGCAAGTATATGCACTTTTTTATAGACTTGAT
                                                         6300
IleValProIleAspAsnAspSerSerThrAsnSerThrAsnTyrArgLeuIleAsnCys
ATAGTACCAATAGACAATGATAGTAGTACCAATAGTACCAATTATAGGTTAATAAATTGT

AsnThrSerAlaIleThrGlnAlaCysProLysValSerPheGluProIleProIleHis
AATACCTCAGCCATTACACAGGCTTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACAT
                                   6400
TyrCysAlaProAlaGlyPheAlaIleLeuLysCysArgAspLysLysPheAsnGlyThr
TATTGTGCCCCAGCTGGTTTTGCGATTCTAAAGTGTAGAGATAAGAAGTTCAATGGAACA

GlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValValSer
GGCCCATGCACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTGGTGTCA
              6500
```

FIG. 7G

```
ThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValIleIleArgSerGluAsn
ACTCAACTGCTGTTGAATGGCAGTCTAGCAGAAGAAGAGGTCATAATTAGATCCGAAAAT
                                                        6600
LeuThrAsnAsnAlaLysAsnIleIleAlaHisLeuAsnGluSerValLysIleThrCys
CTCACAAACAATGCTAAAAACATAATAGCACATCTTAATGAATCTGTAAAAATTACCTGT

AlaArgProTyrGlnAsnThrArgGlnArgThrProIleGlyLeuGlyGlnSerLeuTyr
GCAAGGCCCTATCAAAATACAAGACAAAGAACACCTATAGGACTAGGGCAATCACTCTAT
                                        6700
ThrThrArgSerArgSerIleIleGlyGlnAlaHisCysAsnIleSerArgAlaGlnTrp
ACTACAAGATCAAGATCAATAATAGGACAAGCACATTGTAATATTAGTAGAGCACAATGG

SerLysThrLeuGlnGlnValAlaArgLysLeuGlyThrLeuLeuAsnLysThrIleIle
AGTAAAACTTTACAACAAGTAGCTAGAAAATTAGGAACCCTTCTTAACAAAACAATAATA
            6800
LysPheLysProSerSerGlyGlyAspProGluIleThrThrHisSerPheAsnCysGly
AAGTTTAAACCATCCTCAGGAGGGGACCCAGAAATTACAACACACAGTTTTAATTGTGGA
                                                        6900
GlyGluPhePheTyrCysAsnThrSerGlyLeuPheAsnSerThrTrpAsnIleSerAla
GGGGAATTCTTCTACTGTAATACATCAGGACTGTTTAATAGTACATGGAATATTAGTGCA

TrpAsnAsnIleThrGluSerAsnAsnSerThrAsnThrAsnIleThrLeuGlnCysArg
TGGAATAATATTACAGAGTCAAATAATAGCACAAACACAAACATCACACTCCAATGCAGA
                                        7000
IleLysGlnIleIleLysMetValAlaGlyArgLysAlaIleTyrAlaProProIleGlu
ATAAAACAAATTATAAAGATGGTGGCAGGCAGGAAAGCAATATATGCCCCTCCTATCGAA

ArgAsnIleLeuCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyIle
AGAAACATTCTATGTTCATCAAATATTACAGGGCTACTATTGACAAGAGATGGTGGTATA
            7100
AsnAsnSerThrAsnGluThrPheArgProGlyGlyGlyAspMetArgAspAsnTrpArg
AATAATAGTACTAACGAGACCTTTAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA
                                                        7200
SerGluLeuTyrLysTyrLysValValGlnIleGluProLeuGlyValAlaProThrArg
AGTGAATTATATAAATATAAGGTAGTACAAATTGAACCACTAGGAGTAGCACCCACCAGG

AlaLysArgArgValValGluArgGluLysArgAlaIleGlyLeuGlyAlaMetPheLeu
GCAAAGAGAAGAGTGGTGGAAAGAGAAAAAAGAGCAATAGGATTAGGAGCTATGTTCCTT
                                        7300
GlyPheLeuGlyAlaAlaGlySerThrMetGlyAlaArgSerValThrLeuThrValGln
GGGTTCTTGGGAGCAGCAGGAAGCACGATGGGCGCACGGTCAGTGACGCTGACGGTACAG

AlaArgGlnLeuMetSerGlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGlu
GCCAGACAATTAATGTCTGGTATAGTGCAACAGCAAAACAATTTGCTGAGGGCTATAGAG
            7400
AlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgIle
GCGCAACAGCATCTGTTGCAACTCACGGTCTGGGGCATTAAACAGCTCCAGGCAAGAATC
                                                        7500
LeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
CTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTAGGAATTTGGGGTTGCTCTGGA
```

FIG. 7H

```
LysHisIleCysThrThrAsnValProTrpAsnSerSerTrpSerAsnArgSerLeuAsn
AAACACATTTGCACCACTAATGTGCCCTGGAACTCTAGTTGGAGTAATAGATCTCTAAAT
          .         .         .      7600         .         .
GluIleTrpGlnAsnMetThrTrpMetGluTrpGluArgGluIleAspAsnTyrThrGly
GAGATTTGGCAGAACATGACCTGGATGGAGTGGGAAAGAGAAATTGACAATTACACAGGC
          .         .         .         .         .         .
LeuIleTyrSerLeuIleGluGluSerGlnThrGlnGlnGluLysAsnGluLysGluLeu
TTAATATATAGCTTAATTGAGGAATCGCAGACCCAGCAAGAAAAGAATGAAAAAGAATTG
          .    7700         .         .         .         .
LeuGluLeuAspLysTrpAlaSerLeuTrpAsnTrpPheSerIleThrGlnTrpLeuTrp
TTGGAATTGGACAAGTGGGCAAGTTTGTGGAATTGGTTTAGCATAACACAATGGCTGTGG
          .         .         .         .         .      7800
TyrIleLysIlePheIleMetIleIleGlyGlyLeuIleGlyLeuArgIleValPheAla
TATATAAAAATATTCATAATGATAATAGGAGGCTTGATAGGTTTAAGAATAGTTTTTGCT
          .         .         .         .         .         .
ValLeuSerLeuValAsnArgValArgGlnGlyTyrSerProLeuSerPheGlnThrLeu
GTGCTTTCTTTAGTAAATAGAGTTAGGCAGGGATACTCACCTCTGTCGTTTCAGACCCTC
          .         .         .      7900         .         .
LeuProAlaProArgGlyProAspArgProGluGlyThrGluGluGluGlyGlyGluArg
CTCCCAGCCCCGAGGGGACCCGACAGGCCCGAAGGAACAGAAGAAGAAGGTGGAGAGCGA
          .         .         .         .         .         .
GlyArgAspArgSerValArgLeuLeuAsnGlyPheSerAlaLeuIleTrpAspAspLeu
GGCAGAGACAGATCCGTGAGATTGCTGAACGGATTCTCGGCACTTATCTGGGACGACCTG
          .    8000         .         .         .         .
ArgSerLeuCysLeuPheSerTyrHisArgLeuArgAspLeuIleLeuIleAlaValArg
CGGAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTAATCTTAATTGCAGTGAGG
          .         .         .         .         .      8100
IleValGluLeuLeuGlyArgArgGlyTrpAspIleLeuLysTyrLeuTrpAsnLeuLeu
ATTGTAGAACTTCTGGGACGCAGGGGGTGGGACATCCTCAAATATCTGTGGAATCTCCTA
          .         .         .         .         .         .
GlnTyrTrpSerGlnGluLeuArgAsnSerAlaSerSerLeuPheAspAlaIleAlaIle
CAGTATTGGAGTCAGGAACTGAGGAACAGTGCTAGTAGCTTGTTTGATGCCATAGCAATA
          .         .         .      8200         .         .
AlaValAlaGluGlyThrAspArgValIleGluIleIleGlnArgAlaCysArgAlaVal
GCAGTAGCTGAGGGGACAGATAGAGTTATAGAAATAATACAAAGAGCTTGCAGAGCTGTT
          .         .         .         .  W      .         .
LeuAsnIleProArgArgIleArgGlnGlyLeuGluArgSerLeuLeu         F
                                                       MetGlyGly
CTTAACATACCCAGAAGAATAAGACAGGGCTTAGAAAGGTCTTTACTTTAAAATGGGTGG.
          .    8300         .         .         .         .
 LysTrpSerLysSerSerIleValGlyTrpProAlaIleArgGluArgIleArgArgThr
CAAATGGTCAAAAAGTAGTATAGTGGGATGGCCTGCTATAAGGGAAAGAATAAGAAGAAC
          .         .         .         .         .      8400
 AsnProAlaAlaAspGlyValGlyAlaValSerArgAspLeuGluLysHisGlyAlaIle
TAATCCAGCAGCAGATGGGGTAGGAGCAGTATCTCGAGACCTGGAAAAACATGGGGCAAT
          .         .         .         .         .         .
 ThrSerSerAsnThrAlaSerThrAsnAlaAspCysAlaTrpLeuGluAlaGlnGluGlu
CACAAGTAGCAATACAGCAAGTACTAATGCTGACTGTGCCTGGCTAGAAGCACAAGAAGA
          .         .         .      8500         .         .
 SerAspGluValGlyPheProValArgProGlnValProLeuArgProMetThrTyrLys
GAGCGACGAGGTGGGCTTTCCAGTCAGACCCCAGGTACCTTTAAGACCAATGACTTACAA
          .         .         .         .   U3    .         .
 GluAlaLeuAspLeuSerHisPheLeuLysGluLysGlyGlyLeuGluGlyLeuIleTrp
AGAAGCTCTAGATCTCAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTTG
          .    8600         .         .         .         .
```

FIG. 7I

```
 SerLysLysArgGlnGluIleLeuAspLeuTrpValTyrAsnThrGlnGlyIlePhePro
GTCCAAAAAGAGACAAGAGATCCTTGATCTTTGGGTCTACAACACACAAGGCATCTTCCC
         .         .         .         .         .      8700
 AspTrpGlnAsnTyrThrProGlyProGlyIleArgTyrProLeuThrPheGlyTrpCys
TGATTGGCAAAACTACACACCAGGGCCAGGGATCAGATATCCACTAACCTTTGGATGGTG
         .         .         .         .         .         .
 TyrGluLeuValProValAspProGlnGluValGluGluAspThrGluGlyGluThrAsn
CTACGAGCTAGTACCAGTTGATCCACAGGAGGTAGAAGAAGACACTGAAGGAGAGACCAA
         .         .         .      8800         .         .
 SerLeuLeuHisProIleCysGlnHisGlyMetGluAspProGluArgGlnValLeuLys
CAGCTTGTTACACCCTATATGCCAGCATGGAATGGAGGACCCGGAGAGACAAGTGTTAAA
         .         .         .         .         .         .
 TrpArgPheAsnSerArgLeuAlaPheGluHisLysAlaArgGluMetHisProGluPhe
ATGGAGATTTAACAGCAGACTAGCATTTGAGCACAAGGCCCGAGAGATGCATCCGGAGTT
         .    8900         .         .         .         .
 TyrLysAsn
CTACAAAAACTGATGACACCGAGCTTTCTACAAGGGACTTTCCGCTGGGGACTTTCCAGG
         .         .         .         .         .      9000
GAGGCGTGGACTGGGCGGGACTGGGGAGTGGCTAACCCTCAGATGCTGCATATAAGCAGC
         .   U3    R         .         .         .         .
TGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATTTGAGCCTGGGAGCTCTCTC
         .         .         .      9100         .    B    .
GCTAGCTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAA
         .         .         .         .         .
```